

Simulating feedback and reversibility in substrate-enzyme reactions

D.A.J. van Zwieten¹, J.E. Rooda¹, D. Armbruster^{1,2,a}, and J.D. Nagy³

¹ Department of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands

² School of Mathematical and Statistical Sciences, Arizona State University, Tempe, 85287-1804 AZ, USA

³ Department of Life Sciences, Scottsdale Community College, 9000 E. Chaparral Rd., Scottsdale, 85256 Arizona, USA

Received 22 November 2010 / Received in final form 9 June 2011

Published online 5 August 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract. We extend discrete event models (DEM) of substrate-enzyme reactions to include regulatory feedback and reversible reactions. Steady state as well as transient systems are modeled and validated against ordinary differential equation (ODE) models. The approach is exemplified in a model of the first steps of glycolysis with the most common regulatory mechanisms. We find that in glycolysis, feedback and reversibility together act as a significant damper on the stochastic variations of the intermediate products as well as for the stochastic variation of the transit times. This suggests that these feedbacks have evolved to control both the overall rate of, as well as stochastic fluctuations in, glycolysis.

1 Introduction

In a recent paper Armbruster et al. [1] introduced a novel method of simulating stochastic biochemical networks. Borrowing techniques from production systems, they used discrete events simulation (DES) techniques to perform dynamic simulations of single molecule enzyme networks. It was shown that DES allow the use of non-exponential probability distributions, thus simulating non-Markovian behavior, extending the Gillespie algorithm and its descendants [2] which are typically used to simulate stochastic molecular networks. In this paper we extend the results in [1], which was basically a proof of concept. In particular, we show how discrete event simulations can be used to simulate reversible single substrate-enzyme reactions with positive and negative feedback rules. In this way, gaps in the presentation in [1] are filled, covering the simplest and most common types of feedback present in biochemical pathways.

As an illustration of our basic scheme, Figure 1 shows a simplified version of a substrate enzyme reaction: a substrate molecule S binds with enzyme E , the enzyme re-configures the substrate molecule into product molecule P and the product molecule unbinds from the enzyme



The influence of stochastic variations on the properties of biochemical networks and the interaction of stochasticity and nonlinearity has been an important research topics in recent years. Noise propagation in gene networks [3]

or the influence of noise on the MAPK cascade [4] are prominent topics. The usual approach is to generate a rate equation model for the network, do a steady state analysis of the resulting ODE system and then study the related Langevin equation and the influence of the noise terms on the steady state properties. In some cases, the results are then confirmed by running a Monte-Carlo simulation using a Gillespie algorithm. The fundamental assumption on the stochasticity on all of these studies is that the noise is exponential. The main result of e.g. [4] is that, even though the noise is exponential, the network generates correlations and hence it cannot be treated as modular.

A notable exception to the focus on steady state properties is the study by Rosenfield et al. [5] who determine the response time of a network as a function of the non-linearity in the network. They determine that negative autocorrelation speeds the response time. However, they are not interested in the variations of the response time as a function of the stochastic interaction between the non-linear modules of the network.

Our present paper is not intended to be a significant contribution to a specific bio-chemical problem. However, we want to show that the simulation approach motivated by the similarity of industrial production networks and biochemical networks allows one to extend the current mathematical analysis of biochemical networks in a significant way: we will show that the production-networks inspired approach easily allows for time dependent simulations of the typical nonlinearities encountered in biological reactions and that we do not need the simplification of a Markov assumption in our stochastic processes. The

^a e-mail: armbruster@asu.edu

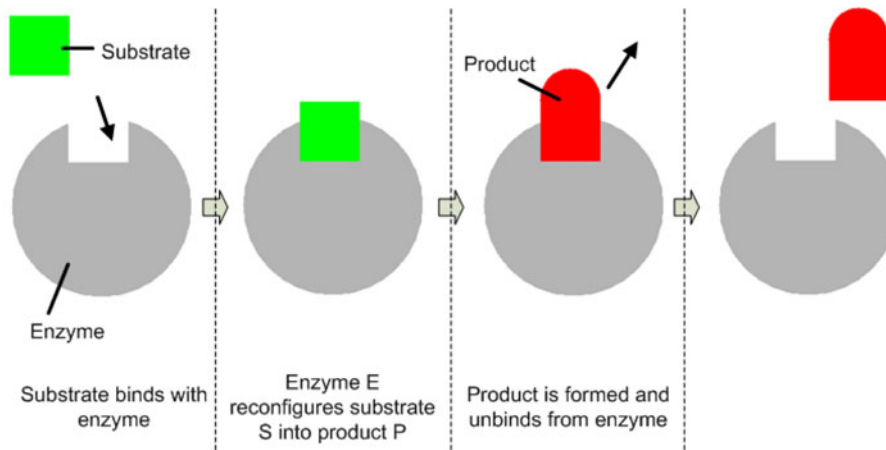


Fig. 1. (Color online) Graphical representation of a substrate-enzyme reaction.

latter seems especially violated for trains of gene regulatory reactions that are spatially localized and are feeding some output of a reaction directly into the next reaction without a stochastic search based on Brownian motion or for ligand-induced conformational motions in motor proteins [6].

As an example, we apply the technique to the Embden-Meyerhof-Parnas (EMP) -glycolysis process. We choose this example for three reasons. First, it is a ubiquitous characteristic of living things. Second, in plants and protists, parts or all of the process are compartmentalized, either by having the controlling enzymes complex with other enzymes or cell structures (like mitochondria in certain plants) or by enclosing glycolytic enzymes within small organelles (e.g., plastids or peroxisomes in plants and Kinetoplastids, respectively) [7]. The most studied example is the Kinetoplastid parasite, *Trypanosoma brucei*, the causative agent of African sleeping sickness, in which the early steps of glycolysis occur in unique organelles called the glycosomes [8] derived from peroxisomes [9]. In such cases, the standard assumptions underlying many in vivo models of glycolysis – large numbers of enzymes in a well-mixed reactor – are unlikely to apply. Third, glycolysis, in particular the enzyme phosphofructokinase (PFK), exhibits a very interesting example of regulation by direct substrates and products. We emphasize, however, that we use this example to illustrate the utility of the technique; the paper is not intended to address specific phenomena associated with glycolysis.

Beyond the application of this simulation technique, we maintain that it is worthwhile to introduce results from operations research/industrial engineering into systems biology. For instance, the understanding that production networks with exponentially distributed service-times have a product form steady state distribution, and hence can be thought of as modular, is only true for a very restricted set of networks (Jackson networks) which is a well known result in queueing theory [10]. The issue of modularity of production networks is a well studied topic [11].

Section 2 gives a short introduction into ODE and discrete event models for substrate-enzyme reactions. The

discrete event model is verified by comparing it with an ODE model. Section 3 presents an expansion of the discrete event model with downstream and/or upstream inhibitory or excitatory feedback affecting the search process or the reconfiguration process. A reversible reaction is modeled and simulated in Section 4. This technique of modeling reversible substrate-enzyme reactions with feedback is then used to model the early steps of the EMP pathway. An analysis of this model is presented in Section 5. Section 6 discusses the influence of stochasticity on the EMP pathway as it occurs in the Kinetoplast parasite, *Trypanosoma brucei*.

2 The substrate-enzyme reaction

2.1 High concentrations – the ODE model

Typically biological models are continuous deterministic models in which variables represent concentrations of species (particles) and their time evolution is governed by ODEs based on average reaction rates. These rates can be measured from experiments looking at the change of concentrations over time. A substrate-enzyme reaction scheme can be separated into the following three reactions:



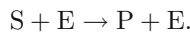
where S denotes the substrate molecule, and E an enzyme molecule. When substrate S and enzyme E interact a complex C is formed. The substrate-enzyme complex can disintegrate into the enzyme E and the substrate S again (2b) or it can fall apart into product P and enzyme E (2c). For high enough concentrations, the stochasticity of the process is assumed to be averaged out, and the

three reactions can be characterized by mass-action rate constants k_1, k_2 and k_3 leading to a 4-dimensional system of mass-action ODEs.

This resulting system has been analyzed in [12–14] and reduced to what is now called the Michaelis-Menten form

$$\frac{d[P]}{dt} = -\frac{d[S]}{dt} = V_{\max} \frac{[S]}{K_m + [S]}, \quad (3)$$

representing the substrate-enzyme reaction scheme



Here the maximal reconfiguration rate V_{\max} is given by $V_{\max} = k_3[E]_t$ with $[E]_t$ the total enzyme concentration and the Michaelis-Menten constant K_m is $K_m = \frac{k_2 + k_3}{k_1}$ [15].

2.2 Low concentrations – the DES model

We review here the discussion in [1]. Simulations of *biochemical networks* consisting of a small number of enzymes and substrate molecules can not rely on the solutions of ODEs but have to take into account the stochastic variations inherent in the individual reactions that no longer average out to a mean behavior. As a result, it is well known that the stochastic trajectory of a biochemical reaction is different from the solution of the mass balance equations. In [1] we apply simulation techniques developed in industrial and production systems engineering to biochemical reaction pathways and show how they may be useful in simulating highly interconnected, nonlinear, nonlocal and non-stationary reaction pathways, which offer no hope of an analytic solution of the underlying evolution equations for the probability densities of individual reactions.

Simulations in manufacturing systems have to manage many different parallel processes. Those processes feed a network of inflows and outflows. At any given time t a number of processes are active. Ending or starting such a process changes the states of the system at separate points in time marking an event. Keeping track of all simultaneous processes and continuously ordering the related events in time is the domain of discrete event simulations. DES in particular have the ability to choose the inter-event times from arbitrarily defined probability distributions, making them a highly versatile tool to perform any kind of stochastic simulations of transport and production processes. The direct connection between manufacturing processes and biochemical reactions is that both processes change an incoming raw material (substrates) into a finished product through repeated interactions with machines (enzymes) and through the addition of parts (ligands) and energy (supplied by ATP).

Typically, randomly fluctuating intracellular biochemical pathways are modeled using the Gillespie algorithm [2], essentially a Markov chain stochastic process in continuous time. The Gillespie algorithm and all its evolved sophistication is based on the simulation of a sample path for the time evolution of the joint probability distribution

for the numbers of molecules of different chemical species that are involved in a particular reaction. Its most important restriction is that it assumes that the reaction is a Markov process and that the time between reactions is exponentially distributed. We will discuss in the next section that this is very questionable and that the reaction time is generated by two independent processes, thus immediately violating the assumption of an exponential distribution for the overall reaction time.

2.3 A single enzyme reaction

A single substrate-enzyme reaction is divided into two distinct processes: the *search* process where the molecules perform a random walk in a predetermined volume, and the *reconfiguration* process where the complex reconfigures the substrate and then disintegrates. In [1] it is argued that the former is a Markov process while the latter is not, a distinction that is easy to implement in a discrete event simulation model: once an enzyme is available for a reaction, the time for the next reaction is given by the search time τ_s which is a random variable. Its realization is determined by sampling from an exponential distribution whose mean depends on the product of the concentrations of the enzymes and the substrate. Similarly, the reconfiguration is a stochastic process. However, that process physically describes an equilibration of energy in the excited molecular complex and hence is better described by a non-Markovian process. We choose a Γ -distribution. Let τ_r be the random variable describing the time for the duration of the reconfiguration process and thus determining the product release. Since for high substrate concentration the search process becomes negligible, τ_r is determined by the asymptote of the Michaelis-Menten curve. Hence $\tau_r = \frac{1}{V_{\max} V}$ where V denotes the volume in which the search process takes place (most likely not the whole cell volume). Simple algebra [1] then gives $\tau_s = \frac{K_m}{[S]} \tau_r$, thus relating the two DES simulation parameters to measurable bulk quantities.

A very similar approach has recently been published by Qian [16]. He develops a theory of single-molecule enzyme kinetics that is time-based and hence focusses on waiting times and cycle times. For instance he finds that the mean waiting time for a single substrate enzyme reaction is the sum of two times: one term that is inversely proportional to the substrate concentration and another one that on average is a constant term. He discusses the statistical properties of this time-based approach for reversible reactions and for feedback systems that lead to co-operative behavior. Our approach agrees completely with Qian's emphasis and complements his theoretical approach by focussing on the simulation aspects that result from this time-based approach. We show below that the production systems perspective naturally leads to Discrete Event Simulations, and we develop the production systems models that are the building blocks of simulations for reversibility and feedback. In addition, since his focus

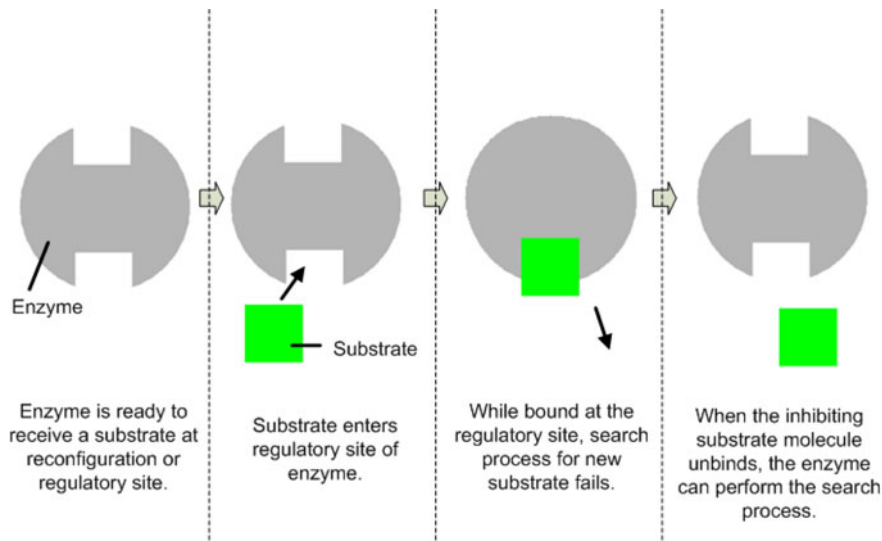


Fig. 2. (Color online) Graphical representation of feedback from a substrate, inhibiting the search for a substrate at the catalytic site. A molecule binding at the feedback site closes the binding site for the catalytic reaction.

was on provable statistics for time-based approaches, Qian had to assume exponential distributions for all stochastic processes and was mostly restricted to equilibrium probabilities and equilibrium relationships. The advantage of DES is that we can study transient phenomena and use any arbitrary stochastic distribution.

3 Feedback

3.1 Feedback impacting the search process

Many substrate-enzyme reactions contain allosteric regulatory feedback loops that depend on the concentration of the substrate or product. The mechanistic understanding of such allosteric feedback assumes that an enzyme molecule has, in addition to its catalytic binding site, regulatory binding sites for a feedback molecule (which may be a substrate or product molecule in our formulation). Binding of a feedback molecule to an enzyme's regulatory site may inhibit or enhance the enzyme's affinity for its substrate. Here we treat the interaction of the allosteric regulator with its target enzyme in a similar manner as we did the interaction between substrate and enzyme. Enzyme-regulator binding is governed by a stochastic search process, and the duration of interaction (length of time enzyme and regulator are bound) is determined by an independent stochastic process, although we do not refer to this as *reconfiguration*, since, in the cases we consider, the regulator is unaltered by the enzyme. Figure 2 presents a graphical representation of substrate inhibition – for example, inhibition of phosphofructokinase by ATP. An inhibiting molecule binding to the enzyme changes the enzyme's structure and prevents any substrate molecules from binding to the catalytic site. If a substrate has already bound to the enzyme and the complex is in the process of reconfiguring itself when a feedback molecule binds, the complex is not affected. However, after the complex

splits into product and enzyme, the search for a new substrate molecule is inhibited until the regulator unbinds.

Such an inhibitory feedback is modeled in a DES by two processes that the enzyme supports, similar to a machine that can produce two different products but only one at a given time. Specifically, once the enzyme is free to bind (i.e. both binding sites are open), two distinct samples – one for the catalytic reaction and one for the allosteric inhibition – are pulled from the distributions of the search times. Let τ_s and τ_{fs} be the search times for the substrate and inhibitor, respectively. Whichever is smaller is taken as the time to the next binding event for that enzyme. Then a second random time is drawn representing the duration of interaction between the enzyme and substrate/inhibitor. Let these interaction times be τ_r and τ_{fr} for the substrate and inhibitor, respectively. If $\tau_{fs} < \tau_s$, then the inhibitor binds first to the free enzyme, which then becomes refractory (has no affinity for its substrate) for the next τ_{fr} time units. If $\tau_s < \tau_{fs}$, then the substrate binds first and a new value for τ_{fs} must be drawn to represent the search process for the inhibitor now with one fewer substrate molecule. (Recall that here we model an enzyme inhibited by its substrate). This resampling is done by sampling again from an exponential distribution since the search process is Markovian.

Results of stochastic simulations of substrate inhibition as depicted in Figure 2 are presented in Figure 3. We choose $v_{\max} = 100.0$, a Michaelis-Menten constant $K_m = 40.0$ and a reconfiguration time $\tau_r = 0.01$. We assume that the substrate molecule binds in the same way to its catalytic and to its regulatory binding site, respectively ($\tau_{fs} = \tau_s = 1$) but has a different release time τ_{fr} at the blocking site. The curve for the deterministic case is generated by setting the variances of the DES to zero. This will result in general in a sawtooth like response curve that is a result of molecular timing issues [17]. The deterministic curve in Figure 3 shows a smooth curve generated when

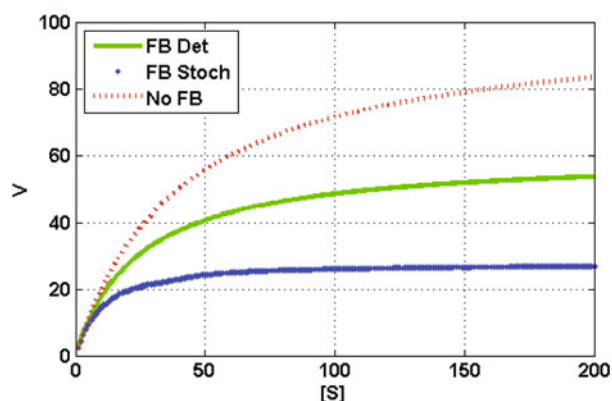


Fig. 3. (Color online) Simulation results for feedback (FB) that inhibits the search process. Catalysis and inhibition have similar kinetics.

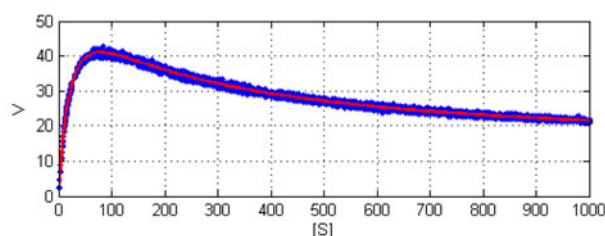


Fig. 4. (Color online) Simulations for inhibition acting on the search process. Vastly different timescales of the catalytic and the inhibitory reaction can generate an atypical response function.

the release time at the blocking site is longer than the re-configuration time. Since the search time is the same for both sites and deterministic, both molecules bind at the same time and the enzyme becomes available again when the blocking is lifted. Adding stochasticity will in 50% of the times lead to a search time for the blocking regulator that is shorter than the search time for the substrate, hence blocking uptake of the substrate by the enzyme. As a result, the stochastic case has about 50% of the out-flux of the deterministic case. The deterministic case will become much more complicated if $\tau_{fr} < \tau_r$. Details are discussed in [18].

In the previous example, since both feedback and catalytic reactions have similar temporal (binding) characteristics, inhibition simply lowers the enzyme's specific activity. Reaction velocity as a function of substrate still shows simple saturation kinetics well described by the Michaelis-Menten equation, but with a lower V_{max} (Fig. 3). However, if inhibition and catalysis have different time scales, the Michaelis-Menten form can be obliterated. For example, if the enzyme's regulatory site has a much lower affinity for the substrate, but retains it longer, once bound, than does its catalytic site, then the reaction velocity as a function of substrate concentration has a unique maximum at a finite substrate concentration (Fig. 4). For this specific case we chose for the substrate molecules $v_{max} = 100$, $K_m = 40$ and for the feedback molecules a binding time of $\tau_{fr} = 0.2$, $K_m = 20\,000$, and a $v_{max} = 1000$.

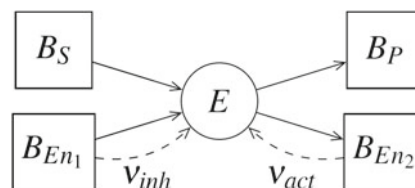


Fig. 5. DEM representation of a process with inhibiting and activating feedback. B_S represents the substrate buffer, B_P the product buffer, B_{En1} and B_{En2} represent the buffers of inhibiting and activating co-substrates, respectively.

Activating feedback can be modeled in a similar way and generates the expected increases in reaction velocities. Since we are not modeling spatial aspects of an enzymatic production line, feedback generated by inhibiting or activating products further upstream or downstream in the line will compete for docking sites at the enzyme in the same way as the immediate product molecule. Hence the simulation aspects of distant regulator molecules are the same as for immediate product regulation, although transients for a distant regulator molecule could be much larger. Often such delays create dynamic instabilities and oscillations, for instance in the MAPK signaling cascade [19].

3.1.1 Inhibiting and activating feedback

The EMP pathway catabolizes glucose to form pyruvate via a multistep biochemical pathway [20]. In addition to the intermediate species derived from glucose, a number of enzymes in this pathway bind ATP or ADP as co-substrates and regulators. The classic example are enzymes generally called phosphofructokinase (PFK), which have catalytic binding sites for fructose 6-P and ATP. PFK catalyzes the transfer of the γ phosphoryl group from ATP to the sugar to form fructose 1,6-bisphosphate and ADP. In addition, many PFK isoforms contain allosteric sites for adenosine. Allosteric binding of ATP tends to inhibit the enzyme, whereas ADP tends to activate it. Figure 5 depicts a general DES model schematic of such a situation. The two additional buffers represent adenylate molecules activating (v_{act}) or inhibiting (v_{inh}) the enzyme (dashed lines).

Figure 6 shows sample paths for simulations of the model diagrammed in Figure 5. In this simulation, activating feedback reduces the search time for a substrate by a factor 50. Inhibition is modeled in the same way as before. Blue and red lines represent substrate and product concentrations, respectively. Dotted lines show sample paths for an equivalent simulation without feedback.

Early in the time course of this simulation the reaction is significantly inhibited by the overwhelming influence of the upstream regulator (the substrate). As the substrate is converted to product, however, the reaction accelerates as inhibition is eased and activation by the product begins to dominate. So, although the early reaction is relatively inefficient compared to the case with no feedback, the reaction in fact runs to completion more rapidly with feedback (Fig. 6).

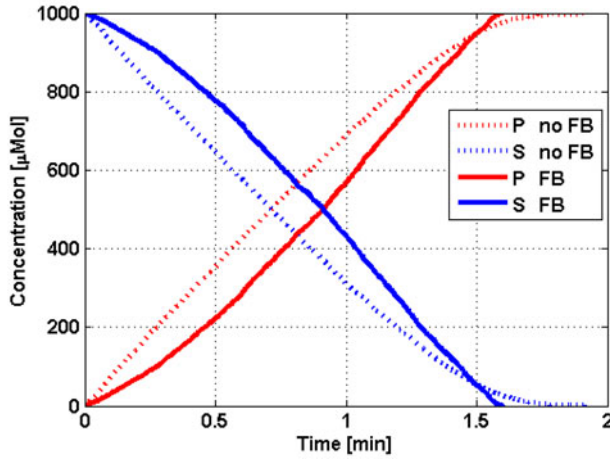


Fig. 6. (Color online) Activating and inhibiting search process feedback simulation results. P and S represent the product concentrations and the substrate concentration, respectively.

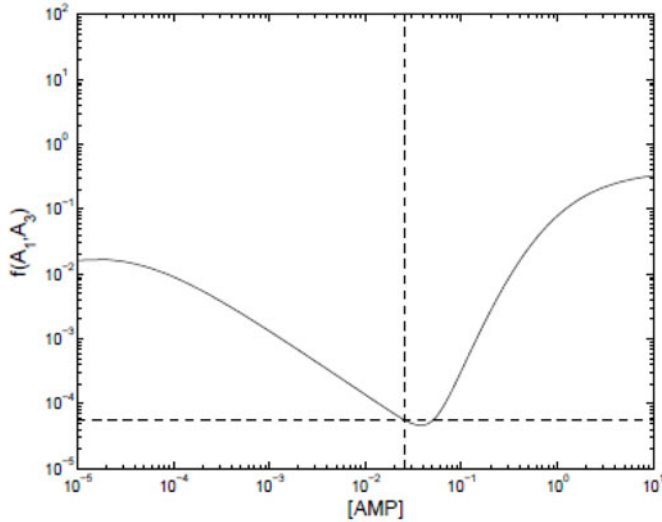


Fig. 7. AMP destruction function $f(A_1, A_3)$ for ATP (A_3) constant and at the default equilibrium value based on a deterministic model. AMP destruction is dominated by 5' nucleotidases when AMP, i.e. A_1 falls below 10^{-1} . Above this threshold, AMP deaminases dominate. Adapted from [21].

3.1.2 Combining enzymes

Total adenylate within a cell is regulated in part by two classes of enzymes that irreversibly degrade AMP: AMP deaminases and 5' nucleotidases. The action of these two classes of enzyme has been modeled assuming complex interaction of an enzyme with both AMP and ATP [21]. In these models, the combined effect of these two enzymes generates a bimodal AMP destruction function, with AMP deaminases and 5' nucleotidases dominating at high and low AMP concentrations, respectively (Fig. 7). In the DES framework, similar situations – that is, bimodal activity curves – can be modeled by combining two parallel enzymes (i.e., parallel machines in a production sys-

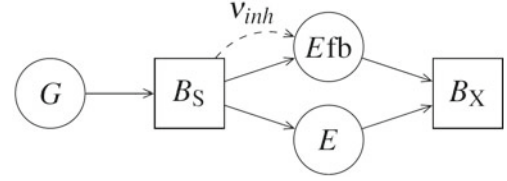


Fig. 8. DEM representation of a combination of a feedback and a normal enzyme. G represents the stochastic generator of substrate molecules, B_S represents the substrate buffer, B_X the product buffer, E represents the regular enzyme and Efb an enzyme that has in addition to its catalytic site a regulatory site that is inhibited by the substrate (represented by the arrow labeled v_{inh}).

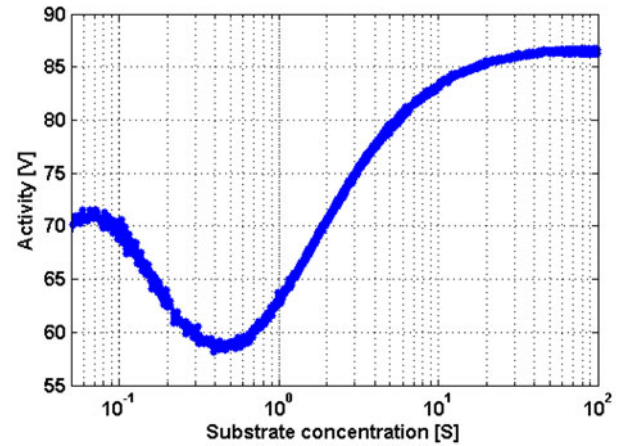


Fig. 9. (Color online) Simulation results for a combination of a feedback and a normal enzyme competing for the substrate with different rates.

tems framework). One enzyme has two binding sites, one catalytic and the other a downstream inhibiting feedback, and the other enzyme has only a single, catalytic site. Both enzymes compete for the same substrate. Figure 8 shows a process structure diagram of this situation, and results of a DES generating a bimodal activity curve are shown in Figure 9.

3.2 Feedback on the reconfiguration process

In some cases, product or substrate molecules interfere with the *reconfiguration process* in an inhibitory or activating manner. The mechanistic assumption is that once a feedback molecule binds at a regulatory binding site, it changes the time evolution of the reconfiguration process without altering the search process (enzyme affinity). Regulation may come from downstream (via the product of the enzymatic reaction) or upstream (via the substrate) or both. A graphical representation of an inhibitory feedback through a product molecule on the reconfiguration process is shown in Figure 10. Since the reconfiguration process is assumed to be non-Markov, the discrete event list has to include the conditional probability for ending the reconfiguration process given that a certain time has already passed since the initial binding. This will happen

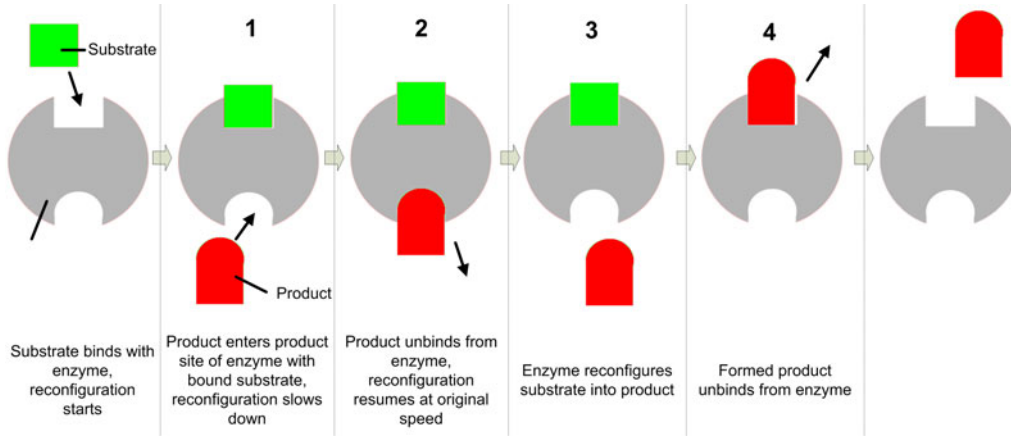


Fig. 10. (Color online) Graphical representation of inhibitory feedback from a product molecule on the reconfiguration process. The labels 1...4 indicate the events discussed in the text.

at the second and fourth step shown in Figure 10. The exact discrete event sequence list is:

- event 1: $t = t_0$. Enzyme finds substrate. Pulls a time $t_r = t_0 + \tau_r$ to release product out of the distribution for the reconfiguration time;
- event 2: $t = t_1 > t_0$. Enzyme finds feedback molecule. Pulls a time $t_{fr} = t_1 + \tau_f$ to release feedback molecule out of the corresponding distribution. Recalculate the reconfiguration time based on the time spent already in reconfiguring the substrate molecule into a product and based on the new distribution influenced by the fact that a feedback molecule is active. This generates a new release time \tilde{t}_r for the product molecule;
- event 3: if $t_{fr} < \tilde{t}_r$ as shown in Figure 10, then at $t_3 = t_{fr}$, the feedback molecule is released and the reconfiguration time is recalculated again based on the two time-periods spent already in reconfiguring the substrate molecule into a product and based on the old distribution without feedback influence. This generates the final release time \bar{t}_r for the product molecule;
- event 4: $t_4 = \bar{t}_r$, the product is released and the enzyme is free to participate in the search for a substrate molecule again.

Other release mechanisms where the product is released before (i.e. $t_{fr} > \bar{t}_r$) or with the feedback molecule ($t_{fr} = \bar{t}_r$) can be devised as necessary.

4 Reversible reactions

Reactions described in the previous sections are all irreversible, i.e. enzyme E only reconfigures substrate S into product P and not the other way around. In reversible reactions the enzyme also reconfigures product molecules P into substrate molecules S. In an ODE model of a reversible reaction the specific activity of the enzyme v_E becomes [22]

$$v_E = \frac{V_{\max} \left([S] - \frac{[P]}{K_{\text{eq},E}} \right)}{K_{m,S} \left(1 + \frac{[P]}{K_{m,P}} \right) + [S]}, \quad (4)$$

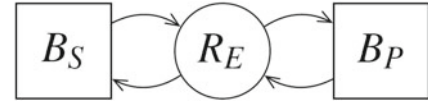


Fig. 11. DEM representation of a reversible reaction.

where $K_{m,x}$ denotes the Henri-Michaelis-Menten constant for x and $K_{\text{eq},E}$ is the equilibrium constant.

The reversible reaction can be modeled as a discrete event model, similar to the substrate-enzyme and feedback model. A DEM representation of the reversible reaction is presented in Figure 11. The blocks B_S and B_P represent the buffers. The buffers contain the substrate and product molecules. The reaction process (enzyme) is presented by R_E .

One way to generate a discrete event simulation is to model the specific activity of the enzyme v_E (4). It can be positive, i.e., reconfiguration of the substrate dominates – or negative – i.e., reconfiguration of the product dominates. When the specific activity is zero an equilibrium has been reached. For a discrete event simulation based on the specific activity (Eq. (4)), the total reaction time Δt (sum of search and reconfiguration times) of one molecule can be calculated from the absolute value of the specific activity of the enzyme $|v_E|$ multiplied by the number of enzymes. The total reaction time for a net increase or decrease by one molecule then becomes:

$$\Delta t = \tau_s + \tau_r = \frac{1}{|v_E| V}. \quad (5)$$

Notice that in equilibrium this time diverges since there is no net change any more. Comparing results of the deterministic DEM with the ODE simulations show a perfect stepwise tracking of the ODE results by the DEM simulations.

Despite these consistent results, however, in reality this approach is physically inaccurate and in particular leads to the wrong stochastic behavior. Specifically, in equilibrium the net flow is zero and hence with a model based on net flows we have no variance of the process. However, in reality, equilibrium is generated by a forward and

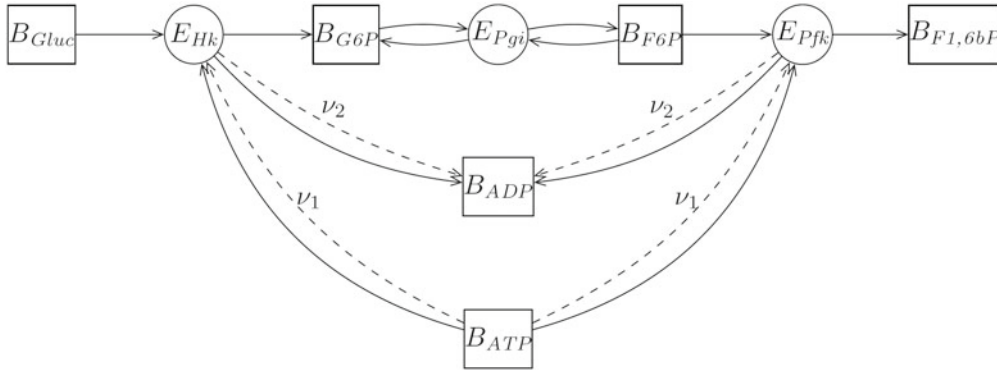


Fig. 12. DEM representation of the EMP-pathway with feedback and reversibility.

backward flow that are equal *on average* and hence, assuming e.g. statistical independence, the variances would add. Hence we model the reversible enzymatic reaction as two competing reactions with search times that are given by the forward specific activity

$$v_F = \frac{V_{\max}[S]}{K_{m,S}\left(1 + \frac{[P]}{K_{m,P}}\right) + [S]}, \quad (6)$$

and backwards specific activity

$$v_B = \frac{V_{\max} \frac{[P]}{K_{eq,E}}}{K_{m,S}\left(1 + \frac{[P]}{K_{m,P}}\right) + [S]}. \quad (7)$$

5 EMP pathway

We illustrate discrete event modeling with feedback and reversibility with a simulation model of compartmentalized glycolysis. Glycolysis is the metabolic pathway that converts glucose into pyruvate. It is highly conserved evolutionarily, and it occurs with variation in essentially all organisms. The most common type of glycolysis is the Embden-Meyerhof-Parnas (EMP) pathway, which was first discovered by Gustav Embden, Otto Meyerhof and Jakub Karol Parnas in 1918. We consider the first steps:

- Phosphorylation of glucose (Gluc) by glucokinase enzymes (Glc) forms glucose-6-phosphate (G6P). In this reaction the γ -phosphoryl group from ATP is transferred to the glucose, generating ADP as a secondary product.
- G6P is then rearranged into fructose-6-phosphate (F6P) by the enzyme phosphoglucosomerase (Pgi). This reaction is freely reversible under normal cell conditions.
- The final reaction we will consider is catalyzed by phosphofructokinase (Pfk). Like Glc, Pfk transfers a γ phosphoryl group from ATP to, in this case, F6P to form the products fructose 1, 6-bisphosphate (F1,6bP) and ADP.

An ODE model of this pathway fragment with kinetic rates is given in, e.g., [22].

The rate of glycolysis is largely regulated by feedback inhibition and activation in this portion of the pathway. The catalytic activity of Pfk in both bacteria and yeast is widely believed to be the key regulator of glycolysis (and therefore much of glucose metabolism) in bacteria, yeast and many other organisms. In its catalytic activity Pfk, in both bacteria and yeast, is inhibited allosterically by ATP. In particular, Pfk has two ATP-binding domains, one catalytic and one regulatory. When ATP is bound at this regulatory site, Pfk takes on a conformation with a relatively low ATP affinity at the catalytic site. In contrast, ADP allosterically activates Pfk. That is, when bound to Pfk's regulatory site, ADP increases the enzyme's catalytic affinity for ATP. These mechanisms are widely believed to provide the main control of the rate of glycolysis. If intracellular ATP concentration falls in the short term, ADP concentration will tend to increase, although this change is buffered somewhat by the action of adenylate kinase, an enzyme that interconverts among the three species of adenosine phosphate. In such a situation, glycolytic throughput increases as Pfk is activated by the increasing ADP concentration. On the other hand, a cell flush with ATP slows glycolytic throughput as Pfk becomes allosterically inhibited by ATP. Upstream feedback activation or downstream inhibition on Glc works similarly. Figure 12 shows the DEM representation of the EMP pathway including reversibility of Pgi and feedbacks ν_1 and ν_2 from ATP and ADP, respectively, on both Pfk and Glc.

Figures 13 and 14 show the influence of reversibility and feedback on the models of the EMP pathway. Figure 13 shows the effects of adding reversibility to the Pgi reaction for a deterministic DES model without feedback. Initial concentrations are $C_{\text{Gluc}} = 1$, $C_{\text{ATP}} = 2$, enzyme concentrations are $0.05 \mu\text{mol}$ and all other initial concentrations are zero. As expected, reversibility greatly slows production of the final product of this pathway fragment, F1,6bP, by reducing the rate at which G6P is converted to F6P, compared to the same model without reversibility.

In our models of feedback regulation, we assume that inhibition of either Glc or Pfk is complete. That is, we assume that when bound to the inhibitor, the enzyme's catalytic site has no affinity for its substrate. Activating feedback, on the other hand, increases the enzyme's affinity.

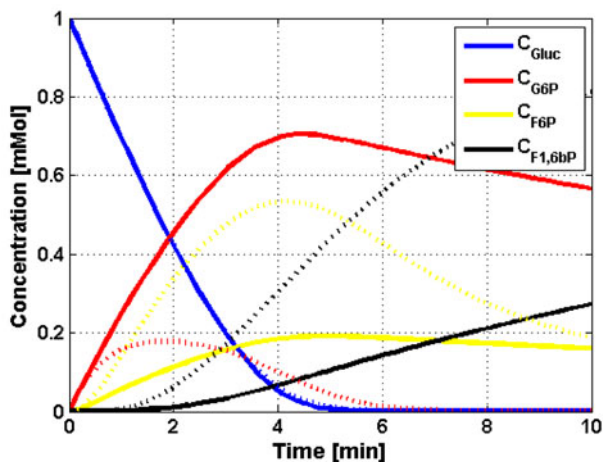


Fig. 13. (Color online) Results without (dotted lines) and with (full lines) reversibility and without feedback. C_x indicates the concentration of molecule x . $C_{\text{Gluc}} = 1$ and $C_{\text{ATP}} = 2$.

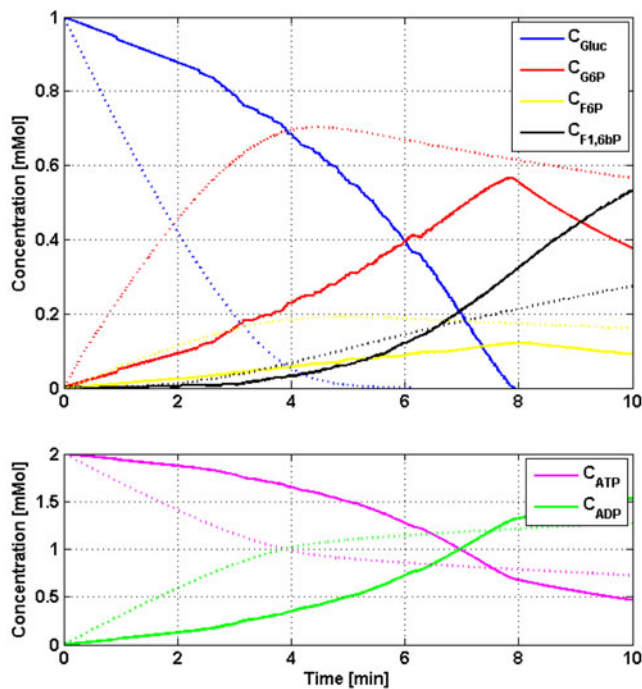


Fig. 14. (Color online) DEM results of the transient simulations of the EMP network with (solid line) and without (dotted line) feedback.

Therefore, we model feedback as affecting the search processes. Feedback search time is calculated similar to the substrate search time. In our DES with feedback, we assume that the binding time of a regulator is on average 1 s. If bound to an inhibitory molecule, the search process is paused until the regulator unbinds. If bound to an activator, we assume that the enzyme's search process is accelerated by a factor of 50. A comparison between simulations of the EMP pathway with and without feedback is presented in Figure 14. Initial concentrations are $C_{\text{Gluc}}(0) = 1.0$ mMol and $C_{\text{ATP}}(0) = 2.0$ mMol. The inhibiting effect of ATP molecules is very large at the begin-

ning and equal to the activating feedback effect at 5 min. After this period activating feedback is dominant. As a result, the throughput with feedback is significantly slowed at first compared to the equivalent model without feedback. However, throughput increases greatly relative to the unregulated case as ATP is depleted and activation overcomes inhibition in Glk and Pfk. This shows that the influence of feedback on this system is significant.

6 Stochastic behavior

The advantage of a discrete event simulation is the ease of including stochastic behavior. While the search process is a random walk in a finite volume and hence the time to a successful search is modeled via an exponential probability distribution, catalysis (the reconfiguration process) is much more complicated. Little is known about the details of the distribution of the reconfiguration times but there is reason to believe that it is not exponential [1]. In order to allow at least the variance as an additional parameter we chose to model the time to product release (τ_r – the reconfiguration time) via a Γ -distribution. We would like to correct an error in [1] here (see also the Erratum [23]). The paper showed the influence of the variance of the Γ -distribution for τ_r on the cycle time, and steady state concentrations of the initial steps (the EPM-steps) of glycolysis. Unfortunately, due to a coding error the coefficient of variation c_v was between 30 and 120, instead of the value $c_v = 3$ as claimed. Obviously that amount of stochasticity is far too high for any biological system. We therefore repeated the simulations for $c_v = 3$ and report them, together with new results on the influence of stochasticity in reversible and feedback reactions in Table 1.

6.1 Variation of the intermediate products

Stochastic behavior of a network of substrate-enzyme reactions leads to stochastic distributions of the characteristic parameters describing the metabolic network like the variations in the levels of intermediate products, or the length of transient timescales. We are especially interested in the difference between the *mean* stochastic behavior and the deterministic simulations based on ODEs. While there is some theory on the qualitative influence of small noise in networks of dynamical systems [24], little is known about the influence of noise in biological networks (see however [25]). To study the influence of noise on the levels of the intermediate products we convert the reversible EMP model with feedback into a steady-state model.

A generator G which represents glucose flux into the glycolytic compartment, an exit process X and a converter process C converting $\text{ADP} \rightarrow \text{ATP}$ are added to the DEM (Fig. 15). The generator G generates substrate molecules S with a certain inter-arrival time t_a . The exit process X represents the final product P (Fru 1,6 bisP) diffusing out of the compartment, and the converter models conversion

Table 1. Comparison between simulations of a steady-state system. We assume a mean interarrival time of glucose molecules of $t_a = 0.025$ and a mean conversion time from ADP back to ATP of $t_c = 0.012$. We list the mean equilibrium levels for Glucose, G6P, F6P, ATP and ADP for a simulation based on deterministic processing rates and stochastic simulations based on exponential search time distributions and Γ -distributed reconfiguration times. The coefficients of variation for the stochastic simulations are also shown.

	Gluc	CV	G6P	CV	F6P	CV	ATP	CV	ADP	CV
Basic										
Det	55.00		26.00		214.00		198.48		0.31	
Stoch	63.48	0.20	26.05	0.19	217.24	0.05	180.19	0.10	18.56	0.93
Difference (%)	15.41		0.19		1.52		-9.22		5830.22	
Rev										
Det	55.00		780.00		214.00		198.48		0.31	
Stoch	63.27	0.19	785.49	0.04	216.95	0.08	180.58	0.09	18.21	0.91
Difference (%)	15.03		0.70		1.38		-9.02		5717.32	
FB										
Det	138.89		25.56		428.96		191.55		7.24	
Stoch	149.47	0.16	26.04	0.19	430.01	0.06	180.97	0.16	17.13	0.92
Difference (%)	7.62		1.88		0.25		-5.52		136.50	
Rev + FB										
Det	138.72		1536.47		429.11		191.64		7.15	
Stoch	139.02	0.09	1543.27	0.03	431.64	0.07	191.37	0.04	7.42	0.94
Difference (%)	0.21		0.44		0.59		-0.14		3.77	

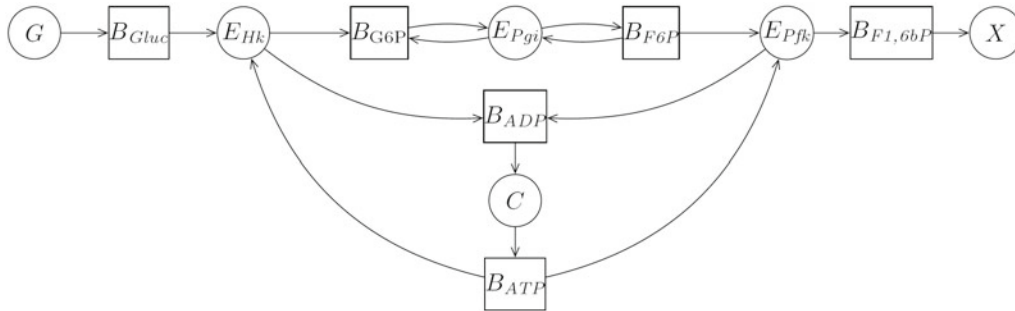


Fig. 15. DEM representation of a steady-state reversible EMP network.

of ADP back to ATP via downstream reactions, which could include the oxidation phase of glycolysis, oxidative phosphorylation, the adenylate kinase reaction and others via an exponential process with an inter-arrival time of t_c .

We emphasize that, as in [1], we are not studying cell metabolism per se, but exploring how a DES approach, here with reversibility and feedback regulation, can be used to study metabolic pathways. In particular, we are not suggesting that adenylate dynamics can be modeled accurately using so simple a system as we employ here.

Table 1 shows a comparison between different simulations of the steady state EMP network. Here we compare deterministic and stochastic models, with all possible combinations of reversibility (with and without) and feedback (with and without). In the stochastic simulations we assume exponential search times and Γ -distributed reconfiguration times, with a $c_v = 3$ as mentioned above. We determine the mean and the c_v values for all molecules. The most interesting observation that can be drawn from

this table is the fact that *feedback and reversibility reduce the stochastic variation of the levels of all substrates involved in glycolysis to almost irrelevant levels*. Note that this is not true for reversibility or feedback alone.

6.2 Variation of the time to decay

Cycle time is a key concept in the study of production systems. It refers to the time that a unit takes to get through a network of machines. Since it is only recently that single molecule reactions can be followed in time for biological systems, the related concept in biology is not yet as important (but see [16]). Cycle time in a manufacturing network run close to capacity can be an order or two of magnitude bigger than the raw production time defined to be the sum of all machine times without any waiting. This difference reflects the effect of congestion in the system. Since there is no identity to a biological molecule and because of the experimental limits discussed above, we study a related

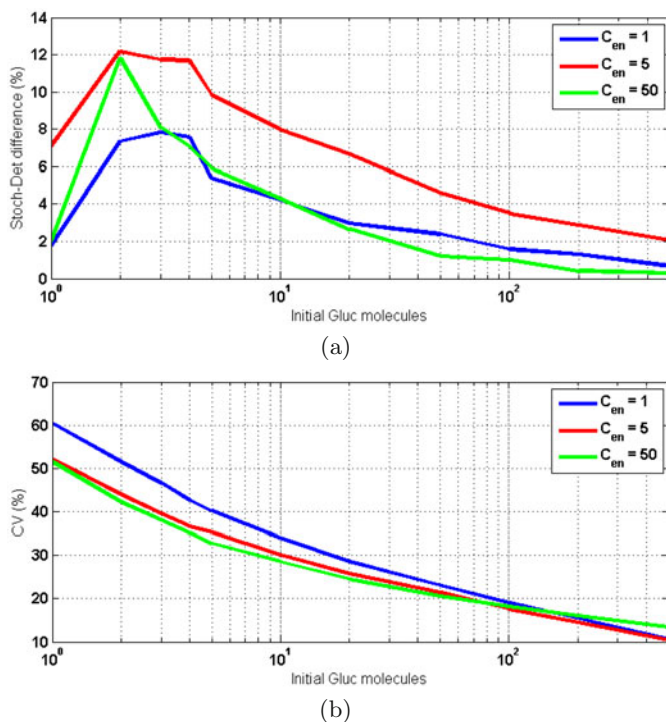


Fig. 16. (Color online) (a) Difference between the mean stochastic and the deterministic transient time for glycolysis. (b) The coefficient of variation for the stochastic simulations. The legend refers to the initial number of enzyme molecules.

concept, the transient time of an excitation through a biological network. In particular, we modeled the whole glycolysis network [18] (12 enzymatic steps) using the DES modeling approach to study the transient decay of one hundred glucose molecules. Figure 16a shows the difference between transient time (defined e.g. to be the time when the number of glucose molecules has dropped to 1% of its initial number) given by the ODE model and the mean stochastic transient for the same system parameters.

We make the somewhat surprising observation that the average transient time of the stochastic process is only about 10% longer than the deterministic transient time. This situation results from two facts: (i) as glucose and its derivatives are depleted over time, congestion in the enzyme buffers becomes less and less important; and (ii) congestion in the enzyme buffers is compensated in part by greater efficiency in the search process when substrate concentrations are high. This latter observation points to a key difference between production lines and biological networks – while an increase in work in progress (WIP) leads to a more or less linear increase in the time to go through a machine, the random walk search process will be more successful when the number of molecules is increased, leading to a *decrease* in the search time and the processing time in an enzyme. Hence, as long as the substrate-enzyme process is dominated by the search process, congestion will not tend to cause a slow-down in the network. However, if the reconfiguration process of the enzyme is much longer

than the average search process, a capacitated machine process results that can generate congestion delays.

7 Conclusion

We have devised schemes to model inhibitory and excitatory feedback on single enzyme reactions via discrete event models and have demonstrated the flexibility of the approach to generate non-standard curves for the reaction rate as a function of the substrate concentrations. In addition we have developed DEM tools to incorporate reversible reactions into these simulations. Discrete event simulations are a natural way to simulate parallel processes, allowing reversible, single enzyme reactions with feedback to be split into different processes corresponding to the different molecular activities in the chemical reaction. In addition, since DEMs are abstracted via the formal methods of process algebras [26,27] simulation programs can be formally verified, which will become more relevant as the programs become more complex.

The main interest for DEMs is their ability to include stochastic behavior into the simulated processes. As a result, DEM simulations not only generate mean behavior but also can easily generate higher moments. Given enough simulation power, the actual probability distributions of any characteristic quantity of a biochemical network can be determined. We show the influence of the stochasticity on the mean and the variance of the number of intermediate products and of the decay time for the EMP steps of glycolysis and highlight the similarities and differences between biological and machine production networks. In most organisms, both eukaryotic and prokaryotic, glycolysis is widely understood to involve large numbers of freely diffusing enzymes and substrates. In bulk (many cell) measurements, stochastic fluctuations will be negligible, and therefore the biochemical dynamics will be close to the mean evolution given by an ODE model. However, it is becoming increasingly clear that such bulk measurements paint a misleading picture of the conditions in individual cells. Evidence is accumulating that cells are not stirred tank reactors. Enzymes are often not free to diffuse, and many metabolic processes are highly compartmentalized, either in specific organelles or even within what has been traditionally understood to be well-mixed cytosol. Even glycolysis, which textbooks always paint as occurring generally throughout the cytosol of eukaryotic cells, is often compartmentalized in plants and protists (reviewed in [7]). Within such compartments, the random walks of metabolic enzymes are very restricted and hence only a small number of molecules will be involved in the reaction network. If that is the case, the variation of the product concentrations and the production time may be high, making a DEM the perfect tool for its simulation. Genetic networks, where a few single mRNA molecules generate protein production [28] is another example for the equivalence of industrial production and biological production.

Although we are not attempting a detailed study of glycolysis here, we note that the DES models suggest that

reversibility and feedback, acting together, function as a significant damper on the stochastic variations of the intermediate products as well as for the stochastic variation of the transit times. In that sense, our analysis, at least for the glycolysis network, justifies the use of mass action ordinary differential equations to study the dynamics in regimes where one would originally expect significant stochastic variation. It is subject to further research to see whether this observation can be generalized.

D.A. was supported by a grant from the Stiftung Volkswagenwerk under the program on Complex Networks and by NSF grant DMS-0604986.

References

1. D. Armbruster, J. Nagy, E.A.F. van de Rijdt, J.E. Rooda, *J. Phys. Chem. B* **113**, 5537 (2009)
2. D.T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976)
3. J.M. Pedraza, A. van Oudenaarden, *Science* **307**, 1965 (2005)
4. S. Tanase-Nicola, P.B. Warren, P.R. ten Wolde, *Phys. Rev. Lett.* **97**, 068102 (2006)
5. N. Rosenfield, M.B. Elowitz, U. Alon, *J. Mol. Biol.* **323**, 785 (2002)
6. Y. Togashi, T. Yanagida, A.S. Mikhailov, *PLoS Comput. Biol.* **6**, e1000814 (2010)
7. M.L. Ginger, G.I. McFadden, P.A.M. Michels, *Phil. Trans. R. Soc. B* **365**, 831 (2010)
8. J.J. Cazzulo, *FASEB J.* **6**, 3153 (1993)
9. J. Moyersoen, J. Choe, E. Fan, W.G. Hol, P.A. Michels, *FEMS Microbiol. Rev.* **28**, 603 (2004)
10. H. Chen, D.D. Yao, *Fundamentals of Queueing Networks: Performance, Asymptotics, and Optimization* (Springer, Verlag, 2001)
11. E. Lefeber, D. Armbruster, Aggregate modeling of manufacturing systems, in *Handbook of Production Planning*, edited by R. Uzsoy, P. Keskinocak, K. Kempf, Kluwer International Series in Operation Research and Management Science (Chapman Hall, 2010)
12. G.E. Briggs, J.B.S. Haldane, *Biochem. J.* **19**, 338 (1925)
13. L. Michaelis, M. Menten, *Biochemische Zeitschrift* **49**, 333 (1913)
14. L.A. Segal, M. Slemrod, *SIAM Rev.* **31**, 446 (1989)
15. S.C. Kou, B.J. Cherayil, W. Min, B.P. English, X.S. Xie, *J. Phys. Chem. B* **109**, 19068 (2005)
16. H. Qian, *Biophys. J.* **95**, 10 (2008)
17. P.W. K hl, M. Jobmann, *J. Recept. Signal Transduction* **26**, 1 (2006)
18. D.A.J. van Zwieten, *Modeling biological networks with discrete event models*, Master's thesis, Eindhoven University of Technology, The Netherlands, 2010
19. L. Qiao, R.B. Nachbar, I.G. Kevrekidis, S.Y. Stanislaw, *PLOS Comput. Biol.* **3**, e184 (2007)
20. B. Teusink, J. Passarge, C.A. Reijenga, E. Esgalhado, C.C. van der Weijden, M. Schepper, M.C. Walsh, B.M. Bakker, K. van Dam, H.V. Westerhoff, J.L. Snoep, *Eur. J. Biochem.* **267**, 5313 (2000)
21. M.V. Martinov, A.G. Plotnikov, V.M. Vitvitsky, F.I. Ataullakhanov, *Biochim. Biophys. Acta* **1474**, 75 (2000)
22. N. Ishii, Y. Sugaa, *FEBS Lett.* **581**, 413 (2007)
23. D. Armbruster, J.D. Nagy, E.A.F. van de Rijdt, J.E. Rooda, *J. Phys. Chem. B* **115**, 7708 (2011)
24. D. Armbruster, E. Stone, V. Kirk, *Chaos* **13**, 71 (2003)
25. A.J. McKane, J.D. Nagy, T.J. Newman, M.O. Stefanini, *J. Stat. Phys.* **128**, 165 (2007)
26. J.C.M. Baeten, D.A. van Beek, J.E. Rooda, Process Algebra, in *Handbook of System Modeling*, edited by P.A. Fishwick (Chapman Hall, Boca Ratan, 2007)
27. D.A. van Beek, K.L. Man, M.A. Reniers, J.E. Rooda, R.R.H. Schiffelers, *The Journal of Logic and Algebraic Programming* **68**, 129 (2006)
28. S. Kar, W.T. Baumann, M.R. Paul, J.J. Tyson, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6471 (2009)

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.